

STIC Biotechnology Systems Branch

RAW SEQUENCE LISTING **ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/683,264
Source: IFW/b
Date Processed by STIC: 1/18/06

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 4.4.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):
U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/10/06

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER:

09/683,264

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor **after** creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line **not exceed** 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do **not** use tab codes between numbers; use **space characters**, instead.
- 4 Non-ASCII The submitted file was **not** saved in ASCII(DOS) text, as **required** by the Sequence Rules. **Please ensure your subsequent submission is saved in ASCII text.**
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. **Per Sequence Rules, each n or Xaa can only represent a single residue.** Please present the **maximum** number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for **each** skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to **include** the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If **intentional**, please insert the following lines for **each** skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000
- 9 Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is **MANDATORY** if n's or Xaa's are present.
 In <220> to <223> section, please explain location of **n** or **Xaa**, and which residue **n** or **Xaa** represents.
- 10 Invalid <213>
 Response Per 1.823 of Sequence Rules, the only **valid** <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is **required** when <213> response is Unknown or is Artificial Sequence
- 11 Use of <220> Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is **MANDATORY** if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n/Xaa "n" can **only** represent a single nucleotide; "Xaa" can **only** represent a single amino acid



IFW16

RAW SEQUENCE LISTING

DATE: 01/18/2006

PATENT APPLICATION: US/09/683,264

TIME: 09:46:35

Input Set : A:\018547.ST25.txt

Output Set: N:\CRF4\01182006\I683264.raw

3 <110> APPLICANT: Affymetrix, Inc.
 4 Siani-Rose, Michael A.
 5 Shigeta, Ron
 7 <120> TITLE OF INVENTION: Computer Software for Automated Annotation of Biological
 8 Sequences
 10 <130> FILE REFERENCE: 018547-048820US
 12 <140> CURRENT APPLICATION NUMBER: US 09/683,264
 13 <141> CURRENT FILING DATE: 2001-12-05
 15 <150> PRIOR APPLICATION NUMBER: US 60/285,144
 16 <151> PRIOR FILING DATE: 2001-04-19
 18 <150> PRIOR APPLICATION NUMBER: US 60/285,403
 19 <151> PRIOR FILING DATE: 2001-04-20
 21 <160> NUMBER OF SEQ ID NOS: 5
 23 <170> SOFTWARE: PatentIn version 3.3
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 373
 27 <212> TYPE: PRT
 28 <213> ORGANISM: Artificial
 30 <220> FEATURE:
 31 <223> OTHER INFORMATION: Figure GRAPAHIT1 sequence
 33 <400> SEQUENCE: 1
 35 Leu Leu Gln Asp Ser Leu Leu Arg Leu Lys Asp Tyr Arg Gln Cys Phe
 36 1 5 10 15
 39 Glu Cys Ser Asp Val Ala Leu Asn Glu Ala Val Gln Gln Met Val Asn
 40 20 25 30
 43 Ser Gly Glu Ala Ala Ala Lys Glu Trp Val Ala Thr Val Thr Gln
 44 35 40 45
 47 Leu Leu Met Gly Ile Glu Gln Ala Leu Ser Ala Asp Ser Ser Gly Ser
 48 50 55 60
 51 Ile Leu Lys Val Ser Ser Ser Thr Thr Gly Leu Val Arg Leu Thr Asn
 52 65 70 75 80
 55 Asn Leu Ile Gln Val Ile Asp Cys Ser Met Ala Val Gln Glu Glu Ala
 56 85 90 95
 59 Lys Glu Pro His Val Ser Ser Val Leu Pro Trp Ile Ile Leu His Arg
 60 100 105 110
 63 Ile Ile Trp Gln Glu Glu Asp Thr Phe His Ser Leu Cys His Gln Gln
 64 115 120 125
 67 Gln Leu Gln Asn Pro Ala Glu Glu Gly Met Ser Glu Thr Pro Met Leu
 68 130 135 140
 71 Pro Ser Ser Leu Met Leu Leu Asn Thr Ala His Glu Tyr Leu Gly Arg
 72 145 150 155 160
 75 Arg Ser Trp Cys Cys Asn Ser Asp Gly Ala Leu Leu Arg Phe Tyr Val
 76 165 170 175

pp 1-4
 Does Not Comply
 Corrected Diskette Needed

*insufficient explanation of
 Artificial Sequence*

*(give source of
 genetic
 material)*

*see item 11 on
 Error Summary
 sheet*

RAW SEQUENCE LISTING

DATE: 01/18/2006

PATENT APPLICATION: US/09/683,264

TIME: 09:46:35

Input Set : A:\018547.ST25.txt

Output Set: N:\CRF4\01182006\I683264.raw

```

79 Arg Val Leu Gln Lys Glu Leu Ala Ala Ser Thr Ser Glu Asp Thr His
80          180          185          190
83 Pro Tyr Lys Glu Glu Leu Glu Thr Ala Leu Glu Gln Cys Phe Tyr Cys
84          195          200          205
87 Leu Tyr Ser Phe Pro Ser Lys Lys Ser Lys Ala Arg Tyr Leu Glu Glu
88          210          215          220
91 His Ser Ala Gln Gln Val Asp Leu Ile Trp Glu Asp Ala Leu Phe Met
92 225          230          235          240
95 Phe Glu Tyr Phe Lys Pro Lys Thr Leu Pro Glu Phe Asp Ser Tyr Lys
96          245          250          255
99 Thr Ser Thr Val Ser Ala Asp Leu Ala Asn Leu Leu Lys Arg Ile Ala
100          260          265          270
103 Thr Ile Val Pro Arg Thr Glu Arg Pro Ala Leu Ser Leu Asp Lys Val
104          275          280          285
107 Ser Ala Tyr Ile Glu Gly Thr Ser Thr Glu Val Pro Cys Leu Pro Glu
108          290          295          300
111 Gly Ala Asp Pro Ser Pro Pro Val Val Asn Glu Leu Tyr Tyr Leu Leu
112 305          310          315          320
115 Ala Asp Tyr His Phe Lys Asn Lys Glu Gln Ser Lys Ala Ile Lys Phe
116          325          330          335
119 Tyr Met His Asp Ile Cys Ile Cys Pro Asn Arg Phe Asp Ser Trp Ala
120          340          345          350
123 Gly Met Ala Leu Ala Arg Ala Ser Arg Ile Gln Asp Lys Leu Asn Ser
124          355          360          365
127 Asn Glu Leu Lys Ser
128          370
131 <210> SEQ ID NO: 2
132 <211> LENGTH: 178
133 <212> TYPE: PRT
134 <213> ORGANISM: Artificial
136 <220> FEATURE:
137 <223> OTHER INFORMATION: Figure 1vhr sequence same env
139 <400> SEQUENCE: 2
141 Ser Val Gln Asp Leu Asn Asp Leu Leu Ser Asp Gly Ser Gly Cys Tyr
142 1          5          10          15
145 Ser Leu Pro Ser Gln Pro Cys Asn Glu Val Thr Pro Arg Ile Tyr Val
146          20          25          30
149 Gly Asn Ala Ser Val Ala Gln Asp Ile Pro Lys Leu Gln Lys Leu Gly
150          35          40          45
153 Ile Thr His Val Leu Asn Ala Ala Glu Gly Arg Ser Phe Met His Val
154          50          55          60
157 Asn Thr Asn Ala Asn Phe Tyr Lys Asp Ser Gly Ile Thr Tyr Leu Gly
158 65          70          75          80
161 Ile Lys Ala Asn Asp Thr Gln Glu Phe Asn Leu Ser Ala Tyr Phe Glu
162          85          90          95
165 Arg Ala Ala Asp Phe Ile Asp Gln Ala Leu Ala Gln Lys Asn Gly Arg
166          100          105          110
169 Val Leu Val His Cys Arg Glu Gly Tyr Ser Arg Ser Pro Thr Leu Val
170          115          120          125

```

RAW SEQUENCE LISTING

DATE: 01/18/2006

PATENT APPLICATION: US/09/683,264

TIME: 09:46:35

Input Set : A:\018547.ST25.txt

Output Set: N:\CRF4\01182006\I683264.raw

173 Ile Ala Tyr Leu Met Met Arg Gln Lys Met Asp Val Lys Ser Ala Leu
 174 130 135 140
 177 Ser Ile Val Arg Gln Asn Arg Glu Ile Gly Pro Asn Asp Gly Phe Leu
 178 145 150 155 160
 181 Ala Gln Leu Cys Gln Leu Asn Asp Arg Leu Ala Lys Glu Gly Lys Leu
 182 165 170 175

185 Lys Pro

189 <210> SEQ ID NO: 3

190 <211> LENGTH: 159

191 <212> TYPE: PRT

192 <213> ORGANISM: Artificial

194 <220> FEATURE:

195 <223> OTHER INFORMATION: Figure 1a17 sequence

197 <400> SEQUENCE: 3

199 Pro Pro Ala Asp Gly Ala Leu Lys Arg Ala Glu Glu Leu Lys Thr Gln
 200 1 5 10 15
 203 Ala Asn Asp Tyr Phe Lys Ala Lys Asp Tyr Glu Asn Ala Ile Lys Phe
 204 20 25 30
 207 Tyr Ser Gln Ala Ile Glu Leu Asn Pro Ser Asn Ala Ile Tyr Tyr Gly
 208 35 40 45
 211 Asn Arg Ser Leu Ala Tyr Leu Arg Thr Glu Cys Tyr Gly Tyr Ala Leu
 212 50 55 60
 215 Gly Asp Ala Thr Arg Ala Ile Glu Leu Asp Lys Lys Tyr Ile Lys Gly
 216 65 70 75 80
 219 Tyr Tyr Arg Arg Ala Ala Ser Asn Met Ala Leu Gly Lys Phe Arg Ala
 220 85 90 95
 223 Ala Leu Arg Asp Tyr Glu Thr Val Val Lys Val Lys Pro His Asp Lys
 224 100 105 110
 227 Asp Ala Lys Met Lys Tyr Gln Glu Cys Asn Lys Ile Val Lys Gln Lys
 228 115 120 125
 231 Ala Phe Glu Arg Ala Ile Ala Gly Asp Glu His Lys Arg Ser Val Val
 232 130 135 140
 235 Asp Ser Leu Asp Ile Glu Ser Met Thr Ile Glu Asp Glu Tyr Ser
 236 145 150 155

239 <210> SEQ ID NO: 4

240 <211> LENGTH: 235

241 <212> TYPE: PRT

242 <213> ORGANISM: Artificial

244 <220> FEATURE:

245 <223> OTHER INFORMATION: Figure GRAPAHIT2 sequence

247 <400> SEQUENCE: 4

249 Pro Leu Cys Lys Gln Ala Leu Glu Asp Leu Glu Lys Thr Ser Gly His
 250 1 5 10 15
 253 Asp His Pro Asp Val Ala Thr Met Leu Asn Ile Leu Ala Leu Val Tyr
 254 20 25 30
 257 Arg Asp Gln Asn Lys Tyr Lys Glu Ala Ala His Leu Leu Asn Asp Ala
 258 35 40 45
 261 Leu Ala Ile Arg Glu Lys Thr Leu Gly Lys Asp His Pro Ala Val Ala
 262 50 55 60

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/683,264

DATE: 01/18/2006

TIME: 09:46:35

Input Set : A:\018547.ST25.txt

Output Set: N:\CRF4\01182006\I683264.raw

```

265 Ala Thr Leu Asn Asn Leu Ala Val Leu Tyr Gly Lys Arg Gly Lys Tyr
266 65          70          75          80
269 Lys Glu Ala Glu Pro Leu Cys Lys Arg Ala Leu Glu Ile Arg Glu Lys
270          85          90          95
273 Val Leu Gly Lys Phe His Pro Asp Val Ala Lys Gln Leu Ser Asn Leu
274          100          105          110
277 Ala Leu Leu Cys Gln Asn Gln Gly Lys Ala Glu Glu Val Glu Tyr Tyr
278          115          120          125
281 Tyr Arg Arg Ala Leu Glu Ile Tyr Ala Thr Arg Leu Gly Pro Asp Asp
282          130          135          140
285 Pro Asn Val Ala Lys Thr Lys Asn Asn Leu Ala Ser Cys Tyr Leu Lys
286 145          150          155          160
289 Gln Gly Lys Tyr Gln Asp Ala Glu Thr Leu Tyr Lys Glu Ile Leu Thr
290          165          170          175
293 Arg Ala His Glu Lys Glu Phe Gly Ser Val Asn Gly Asp Asn Lys Pro
294          180          185          190
297 Ile Trp Met His Ala Glu Glu Arg Glu Glu Ser Lys Asp Lys Arg Arg
298          195          200          205
301 Asp Ser Ala Pro Tyr Gly Glu Tyr Gly Ser Trp Tyr Lys Ala Cys Lys
302          210          215          220
305 Val Asp Ser Pro Thr Val Asn Thr Thr Leu Arg
306 225          230          235
309 <210> SEQ ID NO: 5
310 <211> LENGTH: 233
311 <212> TYPE: PRT
312 <213> ORGANISM: Artificial
314 <220> FEATURE:
315 <223> OTHER INFORMATION: Figure GRAPAHIT3 sequence
317 <400> SEQUENCE: 5
319 Lys Asp Trp Lys Gly Ala Leu Asp Ala Phe Ser Ala Val Gln Asp Pro
320 1          5          10          15
323 His Ser Arg Ile Cys Phe Asn Ile Gly Cys Met Tyr Thr Ile Leu Lys
324          20          25          30
327 Asn Met Thr Glu Ala Glu Lys Ala Phe Thr Arg Ser Ile Asn Arg Asp
328          35          40          45
331 Lys His Leu Ala Val Ala Tyr Phe Gln Arg Gly Met Leu Tyr Tyr Gln
332          50          55          60
335 Thr Glu Lys Tyr Asp Leu Ala Ile Lys Asp Leu Lys Glu Ala Leu Ile
336 65          70          75          80
339 Gln Leu Arg Gly Asn Gln Leu Ile Asp Tyr Lys Ile Leu Gly Leu Gln
340          85          90          95
343 Phe Lys Leu Phe Ala Cys Glu Val Leu Tyr Asn Ile Ala Phe Met Tyr
344          100          105          110
347 Ala Lys Lys Glu Glu Trp Lys Lys Ala Glu Glu Gln Leu Ala Leu Ala
348          115          120          125
351 Thr Ser Met Lys Ser Glu Pro Arg His Ser Lys Ile Asp Lys Ala Met
352          130          135          140
355 Glu Cys Val Trp Lys Gln Lys Leu Tyr Glu Pro Val Val Ile Pro Val
356 145          150          155          160

```

RAW SEQUENCE LISTING

DATE: 01/18/2006

PATENT APPLICATION: US/09/683,264

TIME: 09:46:35

Input Set : A:\018547.ST25.txt

Output Set: N:\CRF4\01182006\I683264.raw

359 Gly Lys Leu Phe Arg Pro Asn Glu Arg Gln Val Ala Gln Leu Ala Lys
360 165 170 175
363 Lys Asp Tyr Leu Gly Lys Ala Thr Val Val Ala Ser Val Val Asp Gln
364 180 185 190
367 Asp Ser Phe Ser Gly Phe Ala Pro Leu Gln Pro Gln Ala Ala Glu Pro
368 195 200 205
371 Pro Pro Arg Pro Lys Thr Pro Glu Ile Phe Arg Ala Leu Glu Gly Glu
372 210 215 220
375 Ala His Arg Val Leu Phe Gly Phe Val
376 225 230

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 01/18/2006
PATENT APPLICATION: US/09/683,264 TIME: 09:46:36

Input Set : A:\018547.ST25.txt
Output Set: N:\CRF4\01182006\I683264.raw

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:1,2,3,4,5

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/683,264

DATE: 01/18/2006

TIME: 09:46:36

Input Set : A:\018547.ST25.txt

Output Set: N:\CRF4\01182006\I683264.raw